

# New pattern of *EGFR* amplification in glioblastoma and the relationship of gene copy number with gene expression profile

Concha Lopez-Gines<sup>1</sup>, Rosario Gil-Benso<sup>1</sup>, Ruben Ferrer-Luna<sup>2</sup>, Rafael Benito<sup>1</sup>, Eva Serna<sup>3</sup>, Jose Gonzalez-Darder<sup>4</sup>, Vicente Quilis<sup>4</sup>, Daniel Monleon<sup>5</sup>, Bernardo Celda<sup>2</sup> and Miguel Cerdá-Nicolas<sup>1</sup>

<sup>1</sup>Department of Pathology, University of Valencia, Valencia, Spain; <sup>2</sup>Department of Quimica-Fisica, University of Valencia, Valencia, Spain; <sup>3</sup>UCIM, University of Valencia, Valencia, Spain; <sup>4</sup>Department of Neurosurgery, Hospital Clínico Universitario, Valencia, Spain and <sup>5</sup>Fundación HCU, Labim, Valencia, Spain

Gene amplification is a process that is characterized by an increase in the copy number of a restricted region in a chromosome arm, and is frequently associated with an overexpression of the corresponding amplified gene. Amplified DNA can be organized either as extrachromosomal elements, repeated units at a single locus or scattered throughout the genome. The amplification of the gene for epidermal growth factor receptor (*EGFR*) is a common finding in glioblastomas and the amplified gene copies appears as double minutes. The aim of this study was to investigate the different patterns of *EGFR* amplification in 40 cases of glioblastoma using FISH analysis in metaphases and paraffin sections, and to investigate the relationship of gene copy number with gene expression profile. The analysis of copy number alterations of *EGFR* was validated by quantitative PCR and SNP microarrays. We observed that in 42% of the cases, the type of amplification of *EGFR* was as double minute chromosomes. In addition, we detected another type of amplification, with extra copies of *EGFR* inserted in different loci of chromosome 7, present in 28% of cases. In this form of amplification, the number of copies is small, and the percentage of cells with *EGFR* amplification is rarely more than 15%. This model of amplification could correspond to a variant of the insertion mechanism, or a consequence of a process of duplication. Our results suggest that this mechanism could represent an early stage of amplification in glioblastomas. Overall, we found a close correlation between *EGFR* gene copy-number alterations and the level of *EGFR* protein expression. However, all cases with a high level of mRNA exhibited strong expression for the *EGFR* protein, and most cases with a low level of mRNA showed no overexpression of *EGFR* protein.

*Modern Pathology* (2010) 23, 856–865; doi:10.1038/modpathol.2010.62; published online 19 March 2010

**Keywords:** glioblastoma; *EGFR* amplification; insertion; microarrays; SNPs

Amplifications are mutations that result in multiple copies of genes in chromosomal regions (amplicons) and induce overexpression in cancer cells. The amount of overexpressed genes in amplified regions varies between different types of cancers, and gene expression, in general, is regulated significantly by DNA copy-number alterations.<sup>1,2</sup> Chromosomal abnormalities associated with DNA amplification

can be organized as extrachromosomal copies, called double minutes (dmns); in tandem arrays as head-to-tail or inverted repeats within a chromosome, often forming a cytologically visible, homogeneously staining region (HSR); or distributed at various locations in the genome (distributed insertions).<sup>1,3,4</sup> The most frequently amplified genes in cancers are *MYCN*, *ERBB2* and *EGFR*.<sup>4</sup>

The epidermal growth factor receptor (*EGFR*) is a cell membrane receptor with intrinsic protein tyrosine kinase activity, which has been the subject of rigorous investigation due to its involvement in several human cancers and its potential as a target of therapy.<sup>5</sup> The activation of *EGFR* is induced by ligand binding that leads to tightly regulated

Correspondence: Professor C Lopez-Gines, Department of Pathology, University of Valencia, Facultad de Medicina y Odontología, Avda. Blasco Ibañez 15, Valencia 46010, Spain.

E-mail: concha.lopez@uv.es

Received 2 November 2009; revised and accepted 28 January 2010; published online 19 March 2010

stimulation of proliferative and survival cell signaling pathways; most notably the RAS/RAF/MEK/MAPK and PI3K/Akt.<sup>6,7</sup>

Glioblastoma multiforme continues to be the most common primary brain tumor in adults with a rapidly progressive and fatal course despite current therapies. The molecular genetics of this tumor is under intense investigation. The *EGFR* gene (7p12.1) is frequently overexpressed, amplified and very rarely mutated in glioblastomas that arise *de novo* (primary glioblastoma).<sup>8</sup>

The first evidence of gene amplification in glioblastoma was provided by cytogenetic analyses that exhibited the presence of *dmin*.<sup>9,10</sup> Molecular screening for gene amplification revealed frequent amplification of the *EGFR*, observed in about 35–70% of glioblastomas.<sup>8,11,12</sup> Differences in the frequency of *EGFR* amplification are most probably due to different methods used, such as Southern blot, polymerase chain reaction (PCR), and fluorescent *in situ* hybridization (FISH).

The overexpression of EGFR protein is usually associated with gene amplification in glioblastomas and these two parameters have been studied as potential prognostic indicators.<sup>13,14</sup> However, it is unclear whether EGFR overexpression would be an accurate prognostic factor in these tumors.

Fluorescent *in situ* hybridization is a powerful methodology to detect *EGFR* copy number abnormalities. This technique has been applied in glioblastomas, in touch preparation and paraffin sections, showing considerable heterogeneity in *EGFR* copy number, although most of the cases considered as amplified presented a high number of copies for this gene.<sup>6,11,14,15</sup> In addition, FISH analysis using *EGFR* probe on tumor metaphases has been applied in a small number of glioblastomas, exhibiting the presence of signals on *dmin*, although HSRs have not been documented in metaphases of glioblastoma.<sup>15,16</sup> Recent introduction of oligonucleotide microarrays designed for whole-genome genotyping of single nucleotide polymorphisms (SNPs) has facilitated the measurement of copy number changes at thousands of SNP loci.<sup>17,18</sup>

We performed FISH analysis using an *EGFR* probe during metaphases of primary cultured cells and in paraffin sections from 40 cases of primary glioblastomas. The aim of this study was to investigate differences in the pattern of *EGFR* amplification in this tumor. We also aimed to compare the *EGFR* copy number alterations, mRNA level and protein expression status, and to correlate the FISH data with the clinical and histopathological parameters in this subset of tumors.

## Materials and methods

### Patient Population and Pathological Study

Tumor samples were obtained from 40 patients with glioblastoma from the Clinic Hospital of Valencia.

All cases were newly diagnosed. None of the patients had received chemotherapy or radiotherapy before surgery. The study was reviewed and approved by Institutional Ethics Committee of the University of Valencia and Clinic Hospital of Valencia. The Karnofsky performance status (KPS) was calculated at the time of diagnosis. All patients underwent surgical removal of the tumor. Subsequently, radiotherapy with 50–65 Gy focal doses was applied to 36 patients. Chemotherapy, including BCNU (bis-chloroethylnitrosourea) or temozolomide, was administered in 31 patients. Survival times were calculated from the day of diagnosis of the glioblastoma to the end date of this study.

The tumor tissue was immediately fixed in neutral-buffered formalin, embedded in paraffin, sectioned and stained with hematoxylin and eosin. The samples were categorized according to the WHO classification<sup>19</sup> and diagnosed as glioblastoma multiforme. Mitotic index values were obtained by counting the total number of mitotic figures in the tumor cells in 10 high-power fields (HPF) and in two different sections. The values are the mean number of mitoses per 10HPF.

The immunohistochemical study was performed on paraffin-embedded sections using the avidin-biotin peroxidase method. The study was carried out using antibodies against glial fibrillary acidic protein (GFAP; Dako, Glostrup, Denmark), Ki-67 (MIB-1, Dako), and monoclonal mouse anti-human *EGFR* (clone H11, Dako), which recognizes the wild-type *EGFR* and the deletion-mutant form of the receptor (*EGFRvIII*).

Proliferation index was evaluated using MIB-1 antibody staining and was calculated by determining the percentage of immunopositive nuclei. The expression of *EGFR* was quantified according to the intensity of staining and number of staining cells, as: 0 (no staining), 1 (light or focal), 2 (moderate), and 3 (strong). Scores of 0 or 1 were defined as no overexpression and scores of 2 and 3 as overexpression.<sup>13</sup>

### Fluorescence *In Situ* Hybridization

To evaluate *EGFR* gene status, dual-color FISH analysis was performed on primary cultured cells and on paraffin sections from 40 samples. The FISH analysis was carried out using the LSR *EGFR* Spectrum Orange/CEP 7 Spectrum Green Probe from Vysis (Abbott Laboratories, Downers Grove, IL, USA; Cat. No. 32-191053).

Cultured cells were treated with Colcemid (0.02 µg/ml; Gibco BRL) for 80 min and chromosomes were prepared by conventional fixation. The paraffin embedded tissues were cut into 5-µm sections and these were mounted on Supoerfrost/Plus microscope slides (Microm International). Hybridizations were performed according to the manufacturer's instructions. Counterstaining of

nuclei was carried out using DAPI. The fluorescent signal was detected using a photomicroscope, Axioplan 2 and Axiophot 2 (Zeiss), equipped with a set of the appropriate filters. Signals were counted in 100–200 cultured cell nuclei, in all the possible metaphases and in 100 non-overlapping tumor cell nuclei in the paraffin sections. The mean signal number for the *EGFR* gene and CEP 7 was calculated for each case, followed by the calculation of *EGFR* gene/CEP 7 ratio. The *EGFR* gene was quantified as amplified in individual cells when the *EGFR*/control signal ratio was greater than 2.<sup>14</sup>

### SNP Array Analysis for Assessment of *EGFR* Copy Number Status

#### *Nucleic acid isolation and microarray experiment*

Fresh frozen tumor tissues (10–20 mg) from 20 glioblastoma patients were used to extract high molecular weight, genomic DNA using DNeasy kit extraction (Qiagen, Valencia, CA, USA). The quality of DNA was assessed by electrophoresis in 1% agarose gel. A total of 250 ng DNA was used for hybridization on Genechip Human Mapping 100 K as recommended by the manufacturer (Affymetrix, Santa Clara, CA, USA). The mapping 100 K set provides a broad coverage of the human genome (92%), with an average distance between SNPs of 23.6 kb. This set includes two arrays, each with more than 50 000 SNPs. The processing was performed following the manufacturer's instructions. Arrays were scanned using GeneChip Scanner 3000. GeneChip operating software was used to define Absent or Present Call and generate CEL files.

#### *Data analysis*

The CEL files were imported into dChipSNP array analysis software.<sup>20</sup> Arrays were normalized against the array with median overall intensity (Baseline array) using the invariant set normalization method.<sup>21</sup> Probe set signal intensity was obtained by using a model based on expression index (PM/MM) method.<sup>22</sup> Average array call rate was >95%. A Hidden-Markov model was used to identify LOH regions from unpaired tumor samples, taking into account SNP intermarker distance, SNP heterozygosity rates, and the haplotype structure of the genome.<sup>23</sup> We used 60 CEPH parents as normal reference genotype and we removed inferred LOH regions consistent with 95% of homozygous markers in normal reference samples. Copy number alterations at each SNP locus were determined using the probe-level signal intensity data.<sup>24</sup> Reference signal distribution was obtained using a trimmed analysis with 80% of the samples.

### Real-Time Quantitative PCR

Tumor tissues from 25 patients with glioblastoma were frozen at  $-80^{\circ}\text{C}$ . Tumor DNA was extracted

from 10–15 mg of the tumor tissue using DNeasy kit extraction (Qiagen). Quantitative PCR (qPCR) was performed using an ABI Prism 7900 HT Fast Real-Time PCR System (Applied Biosystem, Foster City, CA, USA) to analyze *EGFR* copy number. Primer sequences for *EGFR* were: forward: 5'-GTGCAG ATCGCAAAGGTAATCAG-3'; reverse: 5'-GCAGACCGCATGTGAGGAT-3'; and probe: 5'-FAM-CCCCTCCCCGTATCTC-MGB-3'. Primer sequences for *RNase P* and *GAPDH*, used as reference genes, were purchased as TaqMan *RNase P* and *GAPDH* Detection Reagents, respectively (P/N: 4316831).

Target and reference genes were amplified in separate wells. Each 20- $\mu\text{l}$  assay contained 20 ng of genomic DNA, 900 nM each of forward and reverse primers for the gene (*RNase P*, *GAPDH* as reference gene, and *EGFR* as target gene), and 250 nM of labeled gene-specific probe in 1  $\times$  TaqMan Universal PCR Master Mix. The TaqMan universal PCR Master Mix (Applied Biosystems) contained AmpliTaq Gold DNA polymerase, AmpErase uracyl-*N*-glycosylase, deoxynucleotide triphosphates, with dUTPs that replace dTTPs, and optimized buffers.

Individual samples were run in triplicate. The PCR conditions were: 2 min at  $50^{\circ}\text{C}$  (initial incubation for activate AmpErase), 10 min at  $95^{\circ}\text{C}$  (activation of AmpliTaq Gold DNA polymerase) followed by 40 cycles of 15 s at  $95^{\circ}\text{C}$  (melt), and 1 min at  $60^{\circ}\text{C}$  (anneal/extend). Real-time data were collected using the SDS 2.1 software. Each replicate was normalized to reference gene (*RNase P* and *GAPDH*) to obtain  $\Delta C_t$  value, and average  $\Delta C_t$  value for each sample (from the three replicates) was calculated. All samples were normalized to a calibrator sample to determine  $\Delta\Delta C_t$  value. Relative quantity is  $2^{-\Delta\Delta C_t}$  and copy number is  $2 \times$  (relative quantity).<sup>25</sup>

Theoretically, a normal sample has two copies for each gene. Such an assumption is right in cell lines, but taking into consideration that human tumor tissue samples may be contaminated by normal cells, we used a copy-number range approach and considered: homozygous deletion (HD) 0–0.4; borderline HD-loss of heterozygosity (LOH) >0.4–0.6; LOH >0.6–1.4; borderline LOH-normal >1.4–1.6; normal >1.6–2.4; borderline amplification >2.4–2.6; and amplification >2.6.

### *EGFR* mRNA Expression

#### *Nucleic acid isolation*

The RNA was extracted from 19 glioblastoma samples. Five pilocytic astrocytomas without *EGFR* amplification were used as reference. For total RNA extraction, 10–20 mg of frozen tumor samples were homogenized with Ultra Turrax and total RNA was isolated using mirVANA kit (Ambion, Austin, TX, USA) following the manufacturer's instructions. Purified RNA was quantified by UV absorbance at 260 and 280 nm and RNA quality was assessed using a spectrophotometer and Agilent 2100 Bioanalyzer

(Agilent Technologies, Palo Alto, CA USA). Samples with 28S/18S ratio of  $\geq 1.1$  and no evidence of ribosomal peak degradation were included.

#### Microarray experiments

Total RNA (1–15  $\mu\text{g}$ ) was used to generate double-stranded cDNA. cDNA synthesis and cRNA labeling were performed using the protocol for one-cycle cDNA synthesis. Biotin-labeled cRNA (20  $\mu\text{g}$ ) was fragmented and hybridized overnight to Affymetrix HU133plus2.0 GeneChip. Protocols were performed as recommended by Affymetrix. Arrays were washed, stained with streptavidin phycoerythrin, and scanned to generate an image file. Scan quality was assessed by the inspection of visible microarray artifacts, grid placement, background intensity, and housekeeping gene expression. GeneChip operating software (GCOS) was used to define Absent or Present Call and generate CEL files. Arrays with  $< 30\%$  'Present' call for 47 000 probe sets and signal 3'/5' ratio of GADPH control  $0.5 \leq \text{ratio} \leq 4.5$  were omitted.

#### Data analysis

The CEL files were imported into dCHIP program. Arrays were normalized against the array with median overall intensity (baseline array) using the invariant set normalization method.<sup>21</sup> Probe set signal intensity was obtained using a model based on expression index (PM/MM) method.<sup>22</sup> The *EGFR* probe sets with Present Call and expression level  $\geq 10$  in  $\geq 45\%$  samples were selected. Probe sets that passed established criteria were: 1565484\_x\_at, 201983\_s\_at, 201984\_s\_at, 211551\_at, and 211607\_x\_at.

#### Statistical Analysis

In this study, we used no parametric test for Ki67 levels, *EGFR* copy number, *EGFR* transcript expression, and percentage of cells found to be *EGFR* positive by FISH. First, the Kruskal–Wallis test was used to assess significant differences among established groups, and then the comparisons among the pairs of groups were assessed by means of Mann–Whitney *U*-test. We used the Bonferroni method to correct the *P*-values to avoid type-I error.

Pearson's coefficient was used to assess the significant correlation between: *EGFR* copy number measured by qPCR, *EGFR* transcript expression and percentage of *EGFR*-positive cells measured by FISH. The significance of the differential protein expression was contrasted by comparing the *EGFR* amplification type using Fischer's exact test (one tailed). Survival curves for amplification types, KPS groups ( $\leq 80$  and  $> 80$ ), and age ( $< 55$  and  $\geq 55$  years) were evaluated using the Kaplan–Meier method and were compared with the log-rank test. The statistical tests significance was determined at

*P*-value  $\leq 0.05$ . The analysis was performed using SPSS 15.0 software (SPSS, Chicago, IL, USA).

## Results

### Clinical and Histopathological Data

There were 24 male and 16 female patients, reflecting the male predominance among glioblastoma patients. Patient age ranged from 24 to 73 years, with a mean age of 55.5 years. The tumor locations were parietal in 9 patients, temporal in 16, frontal in 10, occipital in 3, and intraventricular in 2 patients. Preoperative functional status was evaluated according to the KPS scale. A total of 30 cases scored 80–100, whereas nine scored 70 and one scored 50. Of the patients, 90% received radiotherapy and 77.5% received chemotherapy. Survival ranged from 3 to 24 months, with a mean survival period of 10.5 months. All patients, except four, were dead when this study was carried out.

Histologically, all tumors demonstrated features of glioblastoma with pleomorphic, astrocytic tumor cells, prominent microvascular proliferation, and necrosis. Five cases showed morphological features of small cell glioblastoma (cases 6, 7, 15, 17, and 24). In each case, the expression of GFAP was confirmed in neoplastic cells. The number of mitoses ranged between 0 and 20, with a mean value of 6. The mean Ki-67 percentage for the tumors was 31.6%; range: 5–80%. The overexpression of *EGFR* (score 2,3) was observed in 20 tumor samples (50%; Table 1).

### *EGFR* Gene and Chromosome 7 Copy Number Alterations

A total of 40 glioblastomas were analyzed successfully by FISH. The FISH results are listed in Table 1. Cells with *EGFR* amplification were seen in 70% of the cases. The metaphases from cultured cells showed: amplification as dmin, in 15 cases (Figure 1a–c), extra copies of *EGFR* inserted in the p or/and q arms of chromosome 7, in 11 cases (Figure 1d–f). Furthermore, two cases showed both types of amplification (Figure 1g). A total of 12 cases of 40 glioblastomas showed no amplification (Figure 1i).

On basis of the *EGFR* status, the gene copy number, and the type of amplification, cases were categorized into three groups: (i) high-level *EGFR* gene amplification and dmin (cases 1–17), (ii) low-level *EGFR* gene amplification and insertions (cases 18–28), and (iii) no *EGFR* gene amplification (cases 19–40). The fraction of amplified cells in group (i) ranged from 1 to 80% in cultured cells (Figure 1a) and from 1 to 92% in cells of paraffin sections (Figure 1j). Only cases 16 and 17 contained less than 10% amplified cells in paraffin sections (Figure 1k). Cells with at least 25 *EGFR* signals were seen in all amplified cases of group (i). There were cells with more than 100 *EGFR* signals in nine cases,

between 50–100 in three cases, and 25–50 signals in three cases. The fraction of amplified cells in group (ii) ranged from 4 to 23% in cultured cells and from 6 to 16% in cells of paraffin sections. These amplified cells showed between 6–25 *EGFR* signals (Figure 1h and l).

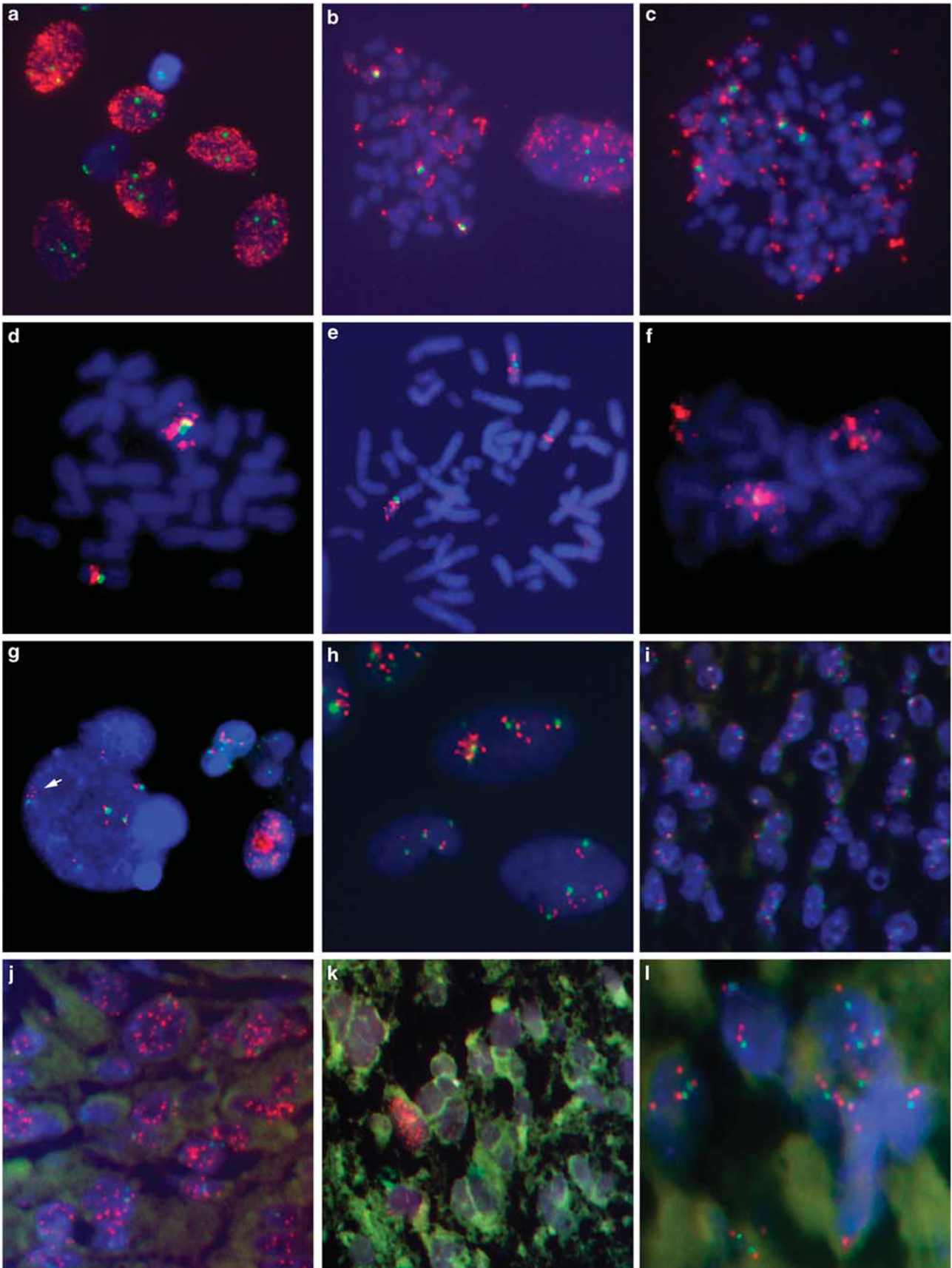
There was considerable heterogeneity in the centromere 7 copy number in these tumors. Amplified cases presented disomy in two, disomy/trisomy in three, trisomy in 12, trisomy/polysomy in nine, and only polysomy in two cases. In total, an excess of chromosome 7 copies was observed in 92% of the

**Table 1** *EGFR* amplification status (FISH, qPCR), chromosome 7, mRNA level, and protein expression in 40 cases of glioblastomas

Case	Cultured cells				Paraffin section			qPCR	mRNA	EGFR (IHC)	MIB-1 %
	Type of amp	EGFR copies/cell	% Of cells	CEP 7	Type of amp	% Of cells	CEP 7				
1	dmin	>100	80	dis	dmin	90	dis	NA	NA	3	40
2	dmin	>100	74	tri/poly	dmin	92	tri/poly	NA	NA	3	10
3	dmin	50–100	21	tri/poly	NA	NA	NA	NA	NA	2	15
4	dmin	>100	57	tri	dmin	70	tri	168.3	1834	3	35
5	dmin	>100	20	tri	dmin	37	tri	55.1	1282	3	25
6	dmin	25–50	65	tri	dmin	73	tri	81.2	1133	3	75
7	dmin	>100	8	tri/poly	dmin	36	tri/poly	81.8	2141	3	50
8	dmin	50–100	6	tri/poly	dmin	80	tri/poly	68.1	2516	2	75
9	NA	NA	NA	NA	dmin	82	tri/poly	77.3	NA	3	10
10	dmin	>100	11	tri/poly	dmin	91	tri/poly	208.2	3125	3	15
11	NA	NA	NA	NA	dmin	51	tri	62.6	3192	3	20
12	dmin	>100	60	tri	no	no	no	97.6	2660	3	10
13	dmin	25–50	75	tri/poly	dmin	73	tri/poly	33.9	2063	3	50
14	dmin	25–50	51	tri	dmin	60	tri	17.8	NA	3	50
15	dmin/ins	50/8–10	80/4	tri	dmin/ins	21	tri	74.6	859	2	75
16	dmin/ins	>100/12–18	9/13	poly	dmin/ins	3/1	poly	NA	NA	1	80
17	dmin	>100	<1	tri	dmin	<1	tri	3.4	144	0	75
18	ins	6–8	15	dis	NA	NA	NA	NA	NA	0	10
19	ins	14–16	4	poly	ins	6	poly	NA	NA	1	50
20	ins	8–12	17	tri	ins	8	tri	NA	NA	1	15
21	ins	6–8	4	tri	NA	NA	NA	NA	NA	0	10
22	ins	6–8	4	dis/tri	NA	NA	NA	NA	NA	0	50
23	ins	6–10	23	dis/tri	ins	10	dis/tri	3.2	120	1	50
24	ins	15–25	9	dis/tri	ins	7	dis/tri	3.9	179	3	5
25	ins	8–10	5	tri	ins	11	tri/poly	3.0	325	0	20
26	ins	10–18	12	tri/poly	ins	13	tri/poly	6.1	141	2	30
27	ins	8–16	23	tri/poly	ins	16	tri/poly	6.5	534	2	10
28	ins	8–10	22	tri	NA	NA	NA	3.7	440	2	20
29	no amp	3–6	—	tri	no amp	—	tri	2.0	NA	0	50
30	no amp	3–10	—	tri/poly	no amp	—	tri/poly	2.0	NA	0	10
31	no amp	3–10	—	tri/poly	no amp	—	tri	NA	NA	1	15
32	no amp	3–10	—	tri/poly	no amp	—	tri/poly	NA	NA	1	75
33	no amp	3–10	—	tri/poly	no amp	—	tri/poly	NA	NA	0	10
34	no amp	3–10	—	tri/poly	no amp	—	tri/poly	2.1	NA	1	25
35	no amp	3–6	—	tri	no amp	—	tri	2.0	NA	0	10
36	NA	NA	NA	NA	no amp	—	tri	2.6	458	0	5
37	no amp	3–10	—	tri/poly	no amp	—	tri/poly	2.3	119	0	25
38	no amp	2–4	—	dis	no amp	—	dis	NA	NA	1	10
39	no amp	3–6	—	tri	no amp	—	tri	NA	NA	2	30
40	no amp	3–6	—	tri	no amp	—	tri	NA	NA	1	20

amp, amplification; CEP 7, centromere chromosome 7; dis, disomy; dmin, doubles minutes; IHC, immunohistochemistry; ins, insertion; tri, trisomy; poly, polysomy; NA, not available; —, no cells with amplification.

**Figure 1** Two different patterns of *EGFR* amplification in glioblastomas. Chromosomal status of this amplification. Probe for *EGFR* and centromere 7 were labeled with red and green respectively. (a) High-level amplification of the *EGFR* in interphase nuclei of cultured cells. (b) Metaphase showing trisomy 7 and *EGFR* amplification as dmin. (c) Polysomy 7 in a metaphase spread with high number of dmin. (d and e) Low-level of amplification: metaphases spread with extra copies of *EGFR* inserted in p and q arms of chromosomes 7 ((d), trisomy 7 with two chromosomes 7 crossed). (f) Multiple extracopies of *EGFR* inserted and distributed in the three chromosomes 7 in a metaphase spread. (g) Different interphase nuclei of cultured cells showing high- and low-level (arrow) of amplification. (h) Low-level of amplification: interphase nuclei showing increased copy number of *EGFR*. (i) Non-amplified tumor cells of paraffin section. (j) Tumor cells showing high-level amplification of the *EGFR* in paraffin section. (k) Isolated amplified tumor cells in paraffin section. (l) Tumor cells in paraffin section with low-level amplification/copy number gains with other cells exhibiting a normal copy number of the *EGFR* gene.



amplified cases. Unamplified cases showed disomy in one, trisomy in four, and trisomy/polysomy in six cases. In this group, 90% cases presented an excess of chromosome 7 copies.

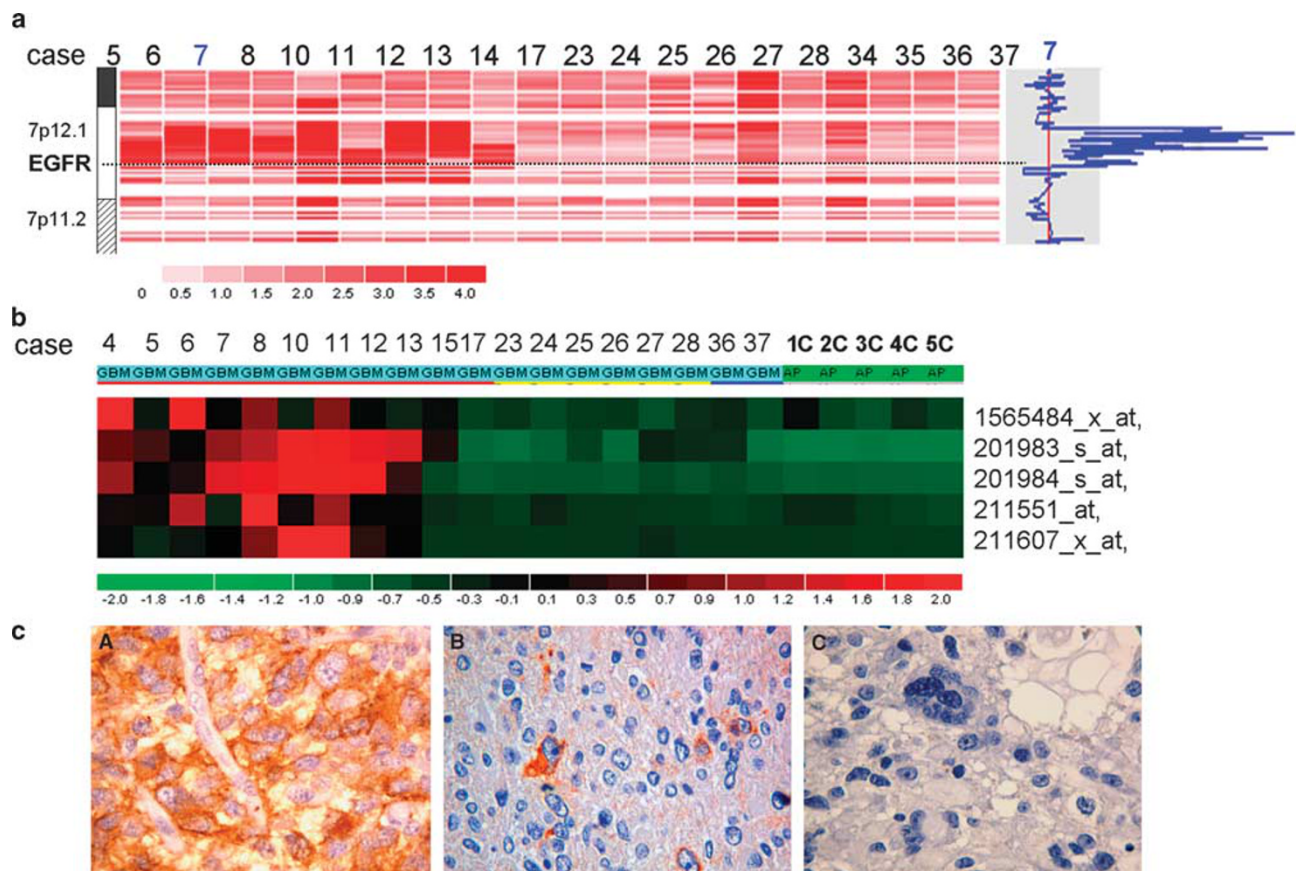
### Correlation of *EGFR* Gene Copy Number, *EGFR* mRNA Expression and EGFR Protein Expression

Using SNP arrays, we inferred the presence of frequent copy-number gains in the region of the *EGFR* locus. In tumor samples, an excess of *EGFR* gene copies was identified by copy number variation from the reference set and from the copy number of the control samples. An increase in the number of *EGFR* copies was seen in tumors of group (i) (Figure 2A). Gene copy numbers at 7p12.1 were validated by qPCR. The qPCR results are listed in Table 1. Each case showed a strong correlation of these values with the copies demonstrated by FISH (Pearson's correlation coefficient = 0.77;  $P$ -value  $\leq 0.01$ ).

Gene expression data were available for 19 glioblastomas (Table 1). We found a significant, positive correlation between the *EGFR* copy number

and the transcript gene expression (Pearson's correlation coefficient = 0.75;  $P$ -value  $\leq 0.01$ ). The increases in *EGFR* dosage were associated with higher levels of *EGFR* transcript (Figure 2B). As expected, the higher percentage of cases of amplified *EGFR* detected by FISH was correlated with a higher gene expression (Pearson's correlation coefficient = 0.85;  $P$ -value  $\leq 0.01$ ).

The EGFR protein expression was evaluated by immunohistochemistry. Cases with a high level of amplification, group (i), manifested EGFR overexpression, ie, scores  $\geq 2$ , except cases 16 and 17. Cases with a low level of amplification, group (ii), had no EGFR overexpression in seven samples and EGFR overexpression in four samples. Cases that showed no amplification of *EGFR*, group (iii), did not express EGFR protein, except case 39 (Figure 2C). Overall, we found a close correlation between *EGFR* gene copy-number alterations and the level of EGFR protein expression in our collection of glioblastomas ( $P$ -value = 0.002; Fisher's exact test). However, all cases with a high level of mRNA exhibited strong expression for the EGFR protein, and most cases with a low level of mRNA showed no overexpression of EGFR protein.



**Figure 2** (a) Heat map integer copy number. Results of 100 K single nucleotide polymorphism (SNP) array analysis in 20 glioblastomas demonstrate genomic gains at 7p12.1, which includes the *EGFR* locus (cases 5–14), which correspond to the (i) group with high-level amplification. (b) *EGFR* gene expression heat map. Red bar denotes tumors with EGFR overexpression and green bar denotes glioblastomas with low expression of this gene. Probe sets are in rows, and samples in columns. (c) Immunoreactivity expression for EGFR: (A) Strong staining, (B) focal staining, and (C) no immunoreactivity. Magnification:  $\times 20$ .

### Statistical Correlation of *EGFR* Gene Copy Alterations with Clinicopathological Parameters

The Kruskal–Wallis test was used to assess the differences among the three amplification groups and it showed significant differences in the percentage of cells with positive *EGFR* amplification detected by FISH, the *EGFR* copy number quantified by PCR, and the *EGFR* transcript expression assessed by microarrays ( $P$ -value  $\leq 0.01$ ). In the percentage of *EGFR*-positive cells and *EGFR* copy number, all bilateral comparisons of the three amplification groups showed statistically significant differences, after the Bonferroni correction (i vs ii, i vs iii, ii vs iii;  $P$ -value  $\leq 0.01$ ). The *EGFR* expression values of amplification types (i) and (ii) were statistically significant after the Bonferroni correction ( $P$ -value  $\leq 0.01$ ). No significant differences were found between types (ii) and (iii). The mean Ki-67 value of group (i) exhibited statistically significant differences with respect the rest ( $P$ -value  $< 0.05$ ). Mean Ki-67 values for different groups were 41.8% for group (i), 24.5% for group (ii), and 23.9% for group (iii).

None of the survival curves showed any significant differences in established groups according to *EGFR* amplification type (i, ii, iii, i + ii vs iii and i vs ii + iii), KPS ( $\leq 80$  and  $> 80$ ), and age ( $< 55$  and  $\geq 55$  years). As expected, younger patients ( $< 55$  years) with KPS  $> 80$  showed a slightly higher survival rate. Interestingly, patients with *EGFR* amplification type (i) survived for the shortest duration; a mean of 8.7 months.

### Discussion

Glioblastoma multiforme displays heterogeneous pathological features associated with complex molecular heterogeneity. Varying genetic alterations lead to the development of malignant phenotype expression and the amplification of *EGFR* is the most frequently involved oncogenetic pathway.

Amplified DNA can be organized as extrachromosomal elements (dmin), as concatenate homogeneously staining chromosomal regions or scattered throughout the genome.<sup>1,3,4</sup> According to the breakage–fusion–bridge model of amplification, the initiating event in HSR formation is double chromatid breakage at a fragile site or telomere erosion,<sup>2,26,27</sup> fused sister chromatids and breaking of the anaphase bridges.<sup>28</sup> Breakage–fusion–bridge cycles could then result in inverted amplified structures,<sup>29</sup> and the mutated sister chromatids are distributed to the daughter cells giving rise to intra-tumor heterogeneity.<sup>30</sup> Gene amplifications are also acquired by selection and unequal segregation of circular extrachromosomal chromatin (dmin and episomes). These elements are formed by looping out from the chromosomes. Dmin and episomes may also relocate in the genome after DNA double-stranded breaks to form HSR or distributed insertions.<sup>2,3,31</sup> Instead of

these above-mentioned mechanisms, N-MYC duplication located at 2p24 has been described in neuroblastoma cell lines, as studied by FISH.<sup>32</sup> Large direct duplications may arise by unequal sister chromatid exchange and seem to have a role during the initial stages in amplification. It is unclear whether duplication represents a prelude to amplification or an alternative pathway for activating the oncogenic potential of *MYC*.<sup>32,33</sup> The single-copy N-MYC is retained at its original site on both, apparently normal, parental copies of chromosome 2 during amplification.

A characteristic cytogenetic feature of glioblastomas is the frequent presence of dmin, found in up to 50% of tumors.<sup>9,10</sup> Molecular studies demonstrated the amplification of several genes in these tumors, especially of the *EGFR* gene; the amplified sequences were found located on dmin, analyzed in a small number of cases by *in situ* hybridization of tumor metaphases.<sup>15,16</sup> However, studies by FISH in interphase nuclei are more frequent, displaying a considerable heterogeneity of *EGFR* copy number.<sup>6,11,14,15</sup> The cases presenting a diffuse rather than a clustered pattern of amplification within single cells support the idea that the most common pattern of *EGFR* amplification in glioblastoma biopsies is dmin. These dmins ranged in size from 0.7 to 2.1 Mb and in level of amplification from 8- to 63-fold.<sup>15</sup> The FISH signals differed from case to case, as intense and multiple, double, faint, and single. This suggests the complexity of dmin formation because both their size and the number of *EGFR* copies they contain, vary. Cell-to-cell variations are also likely to exist for a given tumor.<sup>11</sup>

To the best of our knowledge, this is the first study to analyze the status of the *EGFR* locus at chromosomal level by FISH in a series of 40 newly diagnosed glioblastomas. We observed that in 42% of the cases, the type of amplification is dmin with a variable number of copies affecting over 10% of the cells in most cases. Together with this type of amplification, we have demonstrated another type of amplification present in 28% of the cases, in which extra copies of *EGFR* inserted in different locus of chromosome 7, both in p and q arms, were observed. In this form of amplification, the number of copies is small, and the percentage of cells with the amplification of *EGFR* rarely affects more than 15% of the cells. This model of amplification could correspond to a variant of the distributed insertion mechanism, or a consequence of a process of duplication similar to the one demonstrated in neuroblastoma cell lines.<sup>32</sup> On the basis of these results, we suggest that this mechanism could be considered as an early stage of amplification in glioblastomas. This hypothesis is corroborated by the findings on both status of amplification, dmin, and extra copies, located in chromosome 7, in two cases. In glioblastomas, the cut-off points that define amplification are controversial,<sup>14,34</sup> although the most commonly accepted criterion is that *EGFR*



amplification exists when the *EGFR* gene/chromosome ratio per cell was  $\geq 2$ , or  $\geq 15$  copies of *EGFR* are present per cell in  $\geq 10\%$  of analyzed cells.<sup>35–37</sup>

Microarrays designed for the analysis of SNPs can also be applied in the detection of DNA copy-number alterations. Recent studies have been carried out on glioblastomas and different percentages of *EGFR* amplification have been detected.<sup>7,17,18,38</sup> Our results confirm an excess of *EGFR* copies in cases with a high-level of amplification. The SNP loci with copy-number changes were validated by qPCR. Our cases presented high values of *EGFR* copies, as analyzed by qPCR, in cases with dmin, and low values in cases with insertions. We can confirm a close correlation between these values and those obtained by FISH. However, techniques such as array analysis of SNPs and dPCR (differential) are less likely to detect tumors with isolated *EGFR*-amplified cells or tumors with low-level amplification present in a small subpopulation of tumor cells.<sup>6</sup>

The integration of SNP data with gene expression microarray data is also required to understand the effects of gene regulation and transcription in amplification manifestation. The *EGFR* gene has been found to be expressed at a significantly higher level in glioblastomas.<sup>8,39</sup> The increase in mRNA level caused by amplification is not always proportional to the number of gene copies,<sup>33</sup> and the effects of the gene copy number on expression levels were more relevant for high-level amplifications in a gene-by-gene analysis; results that were also found in breast and prostate cancers.<sup>40,41</sup> In our series, only the cases with dmin amplification, present in more than 10% of the cells, showed an increase of RNA expression.

The *EGFR* protein overexpression is usually associated with gene amplification in glioblastomas.<sup>6,14</sup> In this study, all cases with *EGFR* amplification (group i), except two, showed *EGFR* overexpression; these two cases without *EGFR* overexpression presented less than 10% of amplified cells. In group (ii), only one-third of the cases showed *EGFR* overexpression; and in group (iii), only one case showed *EGFR* overexpression.

In addition, the relationship between trisomy/polysomy of chromosome 7 and *EGFR* gene amplification has been analyzed. In our study, both amplified and non-amplified tumors showed trisomy/polysomy of chromosome 7. This fact suggests that this alteration could be an initial event in the tumorigenesis of glioblastoma.

Amplification in general and amplicons in particular are important for both prognosis and targeted therapies. Specific amplicons, such as *MYCN* in neuroblastomas and *MYC* and *ERBB2* in breast cancer, also have prognostic significance.<sup>4</sup> The *EGFR* amplification status has been studied as a potential prognostic indicator. The majority of the early reports showed that amplification of the *EGFR* was associated with a poorer prognosis than that

associated with non-amplified glioblastomas. Other studies failed to confirm this relationship.<sup>13,14,34,42</sup> Our study did not show that the type of *EGFR* amplification led to significant differences in survival rates; although, interestingly, patients with *EGFR* amplification type (i) survived for the shortest time.

We propose that *EGFR* amplification is an important and frequent pathway in glioblastomas. This amplification is expressed as dmin in a subset of glioblastoma; however, we have observed another type of amplification of the *EGFR* gene. This is seen as distributed insertions of this gene in chromosome 7. Detection of this pattern of amplification of *EGFR* may further improve the prognostic value of genomic diagnosis of this disease.

As amplification is often associated with poor prognosis and is a mechanism of resistance to therapies, it will be important to identify the genes or pathways (or both) that promote amplification in tumors, so that they might be targeted as part of a combination therapy to prevent the evolution of resistance to drugs designed to arrest or eradicate the tumor.

To evaluate the significance of this genomic aberration, it will be necessary to survey a large series of these primary tumors.

## Acknowledgements

We thank Ms Ana Clari for her technical assistance. This study was supported by a grant from FIS, PI061134 (2006–2009), and eTUMOR:FP6-2002-LIFESCIHEALTH 503094.

## Disclosure/conflict of interest

The authors declare no conflict of interest.

## References

- Schwab M. Amplification of oncogenes in human cancer cells. *Bioessays* 1998;20:99–102.
- Mylykangas S, Knuutila S. Manifestation, mechanism and mysteries of gene amplifications. *Cancer Lett* 2006;232:79–89.
- Kuwahara Y, Tanabe C, Ikeuchi T, *et al*. Alternative mechanisms of gene amplification in human cancers. *Genes Chromosomes Cancer* 2004;41:125–132.
- Albertson DG. Gene amplification in cancer. *Trends Genet* 2006;22:447–455.
- Hynes NE, Lane HA. ERBB receptors and cancer: the complexity of targeted inhibitors. *Nat Rev Cancer* 2005;5:341–354.
- Mizoguchi M, Betensky RA, Batchelor TT, *et al*. Activation of STAT3, MAPK, and AKT in malignant astrocytic gliomas: correlation with *EGFR* status, tumor grade, and survival. *J Neuropathol Exp Neurol* 2006;65:1181–1188.
- TCGA (The Cancer Genome Atlas Research Network). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 2008;455:1061–1068.

- 8 Ohgaki H, Kleihues P. Genetic pathways to primary and secondary glioblastoma. *Am J Pathol* 2007;170:1445–1453.
- 9 Bigner SH, Vogelstein B. Cytogenetics and molecular genetics of human brain tumors. *Brain Pathol* 1990;1:12–18.
- 10 Thiel G, Losanowa T, Kintzel D, *et al*. Karyotypes in 90 human gliomas. *Cancer Genet Cytogenet* 1992;58:109–120.
- 11 Sauter G, Maeda T, Waldman FM. Short communication: Patterns of epidermal growth factor receptor amplification in malignant gliomas. *Am J Pathol* 1996;148:1047–1053.
- 12 Okada Y, Hurwitz EE, Esposito JM, *et al*. Selection pressures of TP53 mutation and microenvironmental location influence epidermal growth factor receptor gene amplification in human glioblastomas. *Cancer Res* 2003;63:413–416.
- 13 Shinojima N, Tada K, Shiraishi S, *et al*. Prognostic value of epidermal growth factor receptor in patients with glioblastoma multiforme. *Cancer Res* 2003;63:6962–6970.
- 14 Layfield LJ, Willmore C, Tripp S, *et al*. Epidermal growth factor receptor gene amplification and protein expression in glioblastoma multiforme. *Appl Immunohistochem Mol Morphol* 2006;14:91–96.
- 15 Vogt N, Lefèvre SH, Apiou F, *et al*. Molecular structure of double-minute chromosomes bearing amplified copies of the epidermal growth factor receptor gene in gliomas. *Proc Natl Acad Sci USA* 2004;101:11368–11373.
- 16 Muleris M, Almeida A, Dutrillaux AM, *et al*. Oncogene amplification in human gliomas: a molecular cytogenetic analysis. *Oncogene* 1994;9:2717–2722.
- 17 Wong KK, Tsang YT, Chang YM, *et al*. Genome-wide allelic imbalance analysis of pediatric gliomas by single nucleotide polymorphic allele array. *Cancer Res* 2006;66:11172–11178.
- 18 Wrensch M, McMillan A, Wiencke J, *et al*. Nonsynonymous coding single-nucleotide polymorphisms spanning the genome in relation to glioblastoma survival and age at diagnosis. *Clin Cancer Res* 2007;13:197–205.
- 19 Louis D, Ohgaki H, Wiestler O, *et al*. WHO Classification of Tumors of the Nervous System, 4th edn. IARC: Lyon, 2007.
- 20 Lin M, Wei LJ, Sellers WR, *et al*. dChipPSNP: Significance curve and clustering of SNP-array-based loss-of-heterozygosity data. *Bioinformatics* 2004;20:1233–1240.
- 21 Li C, Wong WH. Model based analysis of oligonucleotide arrays: Expression index computation and outlier detection. *Proc Natl Acad Sci USA* 2001a;98:31–36.
- 22 Li C, Wong WH. Model based analysis of oligonucleotide arrays: model validation, design issues and standard error application. *Genome Biol* 2001b;2:8–14.
- 23 Beroukhi R, Ling M, Park Y, *et al*. Inferring LOH from unpaired tumours using High Density Oligonucleotide SNP arrays. *PLoS Comput Biol* 2006;2:5.
- 24 Zhao X, Li C, Paez G, *et al*. An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. *Cancer Res* 2004;64:3060–3071.
- 25 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta Ct}$  method. *Methods* 2001;25:402–408.
- 26 Coquelle A, Pipiras E, Toledo F, *et al*. Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. *Cell* 1997;89:215–225.
- 27 Murnane JP, Sabatier L. Chromosome rearrangements resulting from telomere dysfunction and their role in cancer. *Bioessays* 2004;26:1164–1174.
- 28 Shimizu N, Shingaki K, Kaneko-Sasaguri Y, *et al*. When, where and how the bridge breaks: anaphase bridge breakage plays a crucial role in gene amplification and HSR generations. *Exp Cell Res* 2005;302:233–243.
- 29 Toledo F, Le Roscouet D, Buttin G, *et al*. Co-amplified markers alternate in megabase long chromosomal inverted repeats and cluster independently in interphase nuclei at early steps of mammalian gene amplification. *Eur Mol Biol Org J* 1992;11:2665–2673.
- 30 Gisselsson D, Pettersson L, Hoglund M, *et al*. Chromosomal breakage–fusion–bridge events cause genetic intratumor heterogeneity. *Proc Natl Acad Sci USA* 2000;97:5357–5362.
- 31 Schwab M. Oncogene amplification in solid tumors. *Cancer Biol* 1999;9:319–325.
- 32 Corvi R, Savelyeva L, Schwab M. Duplication of N-MYC at its resident site 2p24 may be a mechanism of activation alternative to amplification in human neuroblastoma cells. *Cancer Res* 1995;55:3471–3474.
- 33 Schwab M. MYCN in neuronal tumours. *Cancer Lett* 2004;204:179–187.
- 34 Smith JS, Tachibana I, Pohl U, *et al*. PTEN mutation, EGFR amplification, and outcome in patients with anaplastic astrocytoma and glioblastoma multiforme. *J Natl Cancer Inst* 2001;93:1246–1256.
- 35 Cappuzzo F, Hirsch FR, Rossi E. Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. *J Natl Cancer Inst* 2005;97:643–655.
- 36 Hirsch FR, Varella-García M, McCoy J, *et al*. Increased epidermal growth factor receptor gene copy number detected by fluorescence *in situ* hybridization associates with increased sensitivity to gefitinib in patients with bronchioloalveolar carcinoma subtypes: a Southwest Oncology Group Study. *J Clin Oncol* 2005;23:6838–6845.
- 37 Nagasaka T, Gunji M, Hosokai N, *et al*. FISH 1p/19q deletion/imbalance for molecular subclassification of glioblastoma. *Brain Tumor Pathol* 2007;24:1–5.
- 38 Kotliarov Y, Steed ME, Christopher N, *et al*. High-resolution global genomic survey of 178 gliomas reveals novel regions of copy number alteration and allelic imbalances. *Cancer Res* 2006;66:9428–9436. Erratum in: *Cancer Res* 2007; 67:2389.
- 39 Pedersen MW, Pedersen N, Damstrup L, *et al*. Analysis of the epidermal growth factor receptor specific transcriptome: effect of receptor expression level and an activating mutation. *J Cell Biochem* 2005;96:412–427.
- 40 Hyman E, Kauraniemi P, Hautaniemi S, *et al*. Impact of DNA amplification on gene expression patterns in breast cancer. *Cancer Res* 2002;62:6240–6245.
- 41 Wolf M, Mousset S, Hautaniemi S, *et al*. High-resolution analysis of gene copy number alterations in human prostate cancer using CGH on cDNA microarrays: impact of copy number on gene expression. *Neoplasia* 2004;6:240–247.
- 42 Simmons ML, Lamborn KR, Takahashi M, *et al*. Analysis of complex relationships between age, p53, epidermal growth factor receptor, and survival in glioblastoma patients. *Cancer Res* 2001;61:1122–1128.